Journal Name

ARTICLE

Molecular Assembly at Surfaces: Progress and Challenges

R. Raval[a]

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

Molecules provide versatile building blocks, with a vast palette of functionalities and an ability to assemble via supramolecular and covalent bonding to generate remarkably diverse macromolecular systems. This is abundantly displayed by natural systems that have evolved on Earth, which exploit both supramolecular and covalent protocols to create the machinery of life. Importantly, these molecular assemblies deliver functions that are reproducible, adaptable, finessed and responsive. There is now a real need to translate complex molecular systems to surfaces and interfaces in order to engineer 21st century nanotechnology. 'Top-down' and 'bottom-up' approaches, and utilisation of supramolecular and covalent assembly, are currently being used to create a range of molecular architectures and functionalities at surfaces. In parallel, advanced tools developed for interrogating surfaces and interfaces have been deployed to capture the complexities of molecular behaviour at interfaces from the nanoscale to the macroscale, while advances in theoretical modelling are delivering insights into the balance of interactions that determine system behaviour. A few examples are provided here that outline molecular behaviour at surfaces, and the level of complexity that is inherent in such systems.

Introduction

The past two decades have witnessed a step-change in the creation of organised molecular assemblies at surfaces. The scope and versatility of these 2D organic architectures have been captured via scanning probe microscopies, advanced surface spectroscopies and diffraction techniques, as is amply demonstrated by all the contributions to this Faraday Discussion. The variety and abundance of 2D molecular assembly observed at surfaces has illustrated the potential power of 'bottom-up' construction strategies, which combine nanometer-precision and highly parallel fabrication, as a synthesis protocol in new nanotechnologies for sensors, molecular electronics, smart coatings, organic solar cells, catalysts, medical devices, optoelectronics, etc. Such molecule-surface systems are already demonstrating an array of sophisticated functions such as molecular recognition, organisation, responsivity, reactivity, chiral selectivity, energy-harvesting, light manipulation, information storage, and nanolocomotion. The variety of systems, assembly protocols, structures and functions are clearly showcased by researchers contributing to these Proceedings.

However, despite the scientific success that has undoubtedly been achieved, it has to be recognised that the roadmap to creating robust, molecule-surface based devices that can be routinely used for applications is still in its infancy and a number of challenges lie ahead. Chief amongst these is creating a fundamental understanding and control of molecular behaviour and assembly at surfaces and interfaces. Here, the flexibility and adaptability of the molecular state needs to be recognised and tracked. In particular, the manifold of molecule-surface and molecule-molecule interactions need to be captured in order to understand the system from the single molecule perspective to the macroscale collective, from the detailed structure of a single phase to the broader polymorphic phase diagram, from behaviour in controlled conditions to complex environments. This will require the deployment of numerous tools with which to fabricate and interrogate the interface across multiple lengthscales, and close collaborations between experimentalists and theorists. Only then will the generic factors emerge with which to design complex interfaces with intelligent and sophisticated capabilities. Finally, the field will need engineers who can translate this knowledge-base into working devices.

In this Introductory Lecture, I will draw on work that has emerged from my group and collaborations to discuss aspects of molecular behaviour at surfaces, from supramolecular assembly to covalent synthesis. The importance of the nanoscale details in determining global assembly are evident from this work, as is the salutary lesson that, even in

[a] *Surface Science Research Centre and Department of Chemistry, University of Liverpool, Liverpool, L69 3BX, UK. Email: Raval@liverpool.ac.uk*
[b]

controlled environments, the molecule-surface system invariably exhibits complex behaviour.

Highly Organised Molecular Assembly at Surfaces

The behaviour of molecules at defined single crystal metal surfaces, HOPG, oxide layers, etc. has been instrumental in demonstrating the organisational ability of molecules at interfaces, displaying highly ordered assemblies with precise molecular arrangements that can propagate from the nanoscale to the macroscale [1-6]. Generally, these organised assemblies fall into two main groups: (i) those which involve weak interactions with the surface, with supramolecular interactions dominating system behaviour; and, (ii) those that involve strong bonding to the surface, with the resulting behaviour determined by a balance of molecule-surface, molecule-molecule and intra-surface interactions. I will concentrate on the latter systems, but point here to the excellent examples of both types of systems by other contributions in this Faraday Discussion. Both sets of systems show aspects that are challenging to the experimentalist and the theorist. Foremost, a significant proportion of the systems display polymorphism where the nature of the molecular entity and its organisation vary significantly as surface coverage, temperature and deposition flux are varied [2]. These aspects often remain unexplored, with researchers focussing on structure-determination of specific organised phases only. Even when studies are targeted at a single phase populated by just one type of molecular species, many structural models can often be suggested, with only slight variations in energy. Such aspects constitute a central debate in 3D molecular crystals [7] and these considerations are now beginning to emerge in the study of 2D systems. Two examples described below on the assembly of tartaric acid and the amino acid, proline, on the Cu(110) surface illustrate the complex nature of supramolecular assembly at surfaces, especially when the molecules have a strong chemisorption interaction with the surface. In addition, both molecules are chiral, adding further complexity arising from the nature of enantiomeric interactions and their self-assembly behaviour.

A second advance in the field is the creation of covalent macromolecular structures at surfaces, bringing the domain of organic covalent synthesis to the surface and the interface. This is a very important development in the field, with rapid progress since the first reports emerged [8-11]. I will again draw on our work on dehydrogenative coupling via C-H activation and scission to illustrate some of the state-of-the-art in the field. Finally, the importance of kinetics and transport in dictating the assembly that emerges under particular conditions surely remains as an outstanding contemporary challenges, and will require development of both experimental and theoretical approaches.

Chiral Supramolecular Assembly of Tartaric Acid on Cu(110): some recent developments.

The assembly of chiral molecules at surfaces has become an important area in its own right, with surface spectroscopies and scanning probe microscopy delivering important insights on chiral ordering, recognition, heterogeneous enantioselectivity and mirror-symmetry breaking to produce organised homochiral assemblies [2,3,12-33].

An exemplar in the field of surface chirality is provided by the tartaric acid/Cu(110) system [12-18]. This system displays a high degree of polymorphism with the nature of the molecular building block dynamically changing between the intact bi-acid, the singly dehydrogenated monotartrate and the doubly dehydrogenated bitartrate, as revealed by a combination of Reflection Absorption Infrared Spectroscopy (RAIRS), Temperature Programmed Desorption (TPD), Low Energy Electron Diffraction (LEED) and Scanning Tunnelling Microscopy (STM) experiments [2,13,22]. One phase, in particular, created by the bitartrate species has attracted attention because it displays global space group homochirality for the enantiopure system [12], with *(R,R)*-bitartrate assembling into the (1 2, -9 0) structure and *(S,S)*-bitartrate into the mirror (9 0, -1 2) assembly, Figure 1a,b, with hierarchical transfer of homochirality from the single-molecule to the macroscale assembly. The racemic system displays Pasteurian chiral resolution at the surface, Figure 1c, with enantiomers segregating into mirror domains of (1 2, -9 0) and (9 0, -1 2) structure. Recently, we re-investigated this system [18] using STM and dispersion corrected periodic DFT calculations that are now able to handle the large unit cell of the system. This work showed that the original calculated structure [14, 15] for this system was incorrect, unable to either account for the intensity variation observed in the STM images, or the chiral segregation observed for the racemic system. This work also revealed the level of detailed analysis required to understand complex chiral molecule assembly at surfaces.

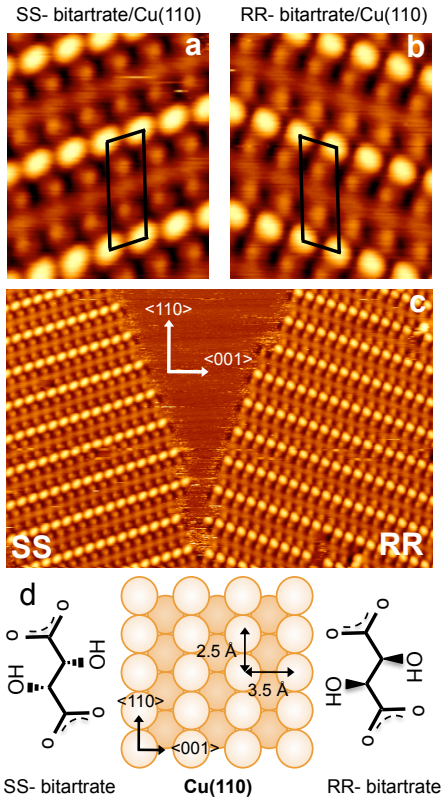


Figure 1. STM images of bitartrate on Cu(110). Top panel: STM image of the global homochiral assembly of a) *(S,S)*-bitartrate on Cu(110), (45 x 52 Å^2, I_t = 0.23nA, V_t = 0.541V) and b) *(R,R)*-bitartrate on Cu(110) (45 x 52 Å^2, I_t = 0.23nA, V_t = 0.541V.) c) STM image of the chiral segregation observed after adsorption of racemic tartaric acid on Cu(110), showing domains of (9 0, -1 2) *(S,S)* and (1 2, -9 0) *(R,R)* bitartrate (277 x 193 Å^2, I_t = 0.21nA, V_t = 0.557V). d) Structures of the two enantiomers of bitartrate and the Cu(110) surface.

The starting point for analysing molecular assembly at a surface is to understand the nature of the adsorbed molecular species. Calculations for the enantiopure system show that *(R,R)*-bitartrate chemisorbs strongly to the Cu(110) surface by bridging across two close packed Cu rows to create a 4-point bonded adsorption motif, with both carboxylate groups bonded to the surface by four Cu-O bonds, Figure 2. Interestingly, it is the detail of the adsorption footprint that holds the key to system behaviour, specifically the chiral segregation that is observed for the racemic mixture. Three distinct adsorption motifs are possible, one transcribing a rectangular footprint (R_{ec}) and two producing oblique footprints (O_a or O_b), which are mirror images. Each change in footprint leads to significant local changes in the adsorbed molecules. Adsorption in the rectangular footprint leads to strong hydrogen bonds between the α–hydroxyls and the adjacent carboxylate oxygen, causing a chiral distortion in the molecular backbone, with a C-C torsion angle of -28°. This chiral distortion is increased further with a C-C torsion angle of -50° if the O_a oblique footprint is adopted. Concomitantly, there is significant restructuring of the internal hydrogen bonds. This adsorption motif is just 0.083 eV less stable than the rectangular footprint model. If *(R,R)*-bitartrate adopts the mirror O_b oblique footprint, the carbon skeleton distorts in the opposite direction with a C-C torsion angle of +41°, the internal hydrogen bonds are weakened, and the O_b structure is 0.42 eV less stable than the rectangular R_{ec} footprint. This is a high energy penalty, so this *(R,R)*-bitartrate conformation can be excluded from consideration for self-assembly. However, the R_{ec} and O_a footprints are both candidates for 2D assembly of *(R,R)*-bitartrate into the experimentally observed (1 2, -9 0) structure. Therefore, all allowed permutations of the R_{ec} and O_a conformations need to be considered and, in order to discriminate between the various structural models, three major questions need to be addressed:

i. Is the assembly structure energetically preferred?
ii. Does the structure predict the chiral segregation observed for the racemic mixture?
iii. Does the structure produce the observed STM image?

Experimental observations show that the enantiopure (1 2, -9 0) *(R,R)*-bitartrate structure possesses rows of bitartrate trimers that assemble with high fidelity across the entire surface. Each trimer is aligned along a chiral direction, and

each row of trimers also assembles in a chiral direction with respect to the surface, as shown in Figures 1 and 3 (top row). Four distinct assemblies using the R_{ec} and the O_a footprints need to be considered, shown as Structures 1 – 4 in Figure 3. The inserted table in Figure 3 lists the computed energies of the modelled structures [18]. Structure 4 with a $O_aO_aO_a$ footprint combination can be excluded since it has a significantly higher energy compared to the other three. Structure 2 has the lowest energy, with 1 and 3 only slightly higher. Therefore, all three are viable candidates on energetic grounds.

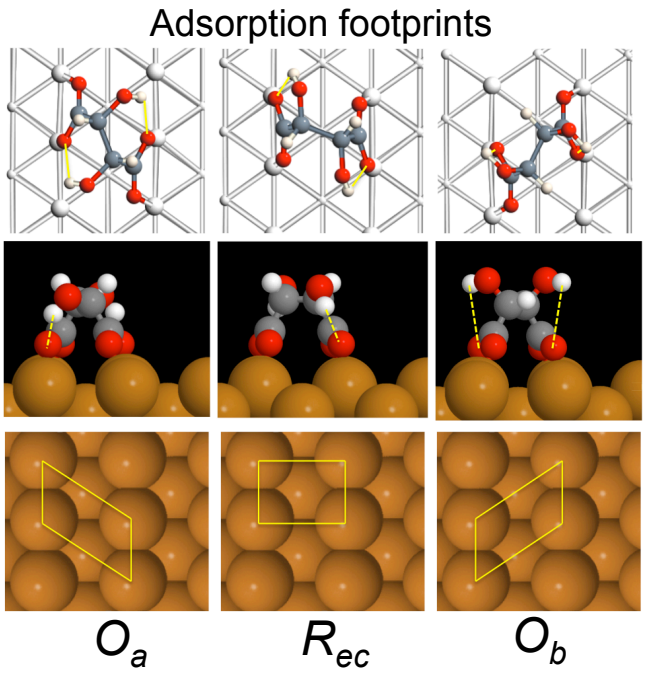
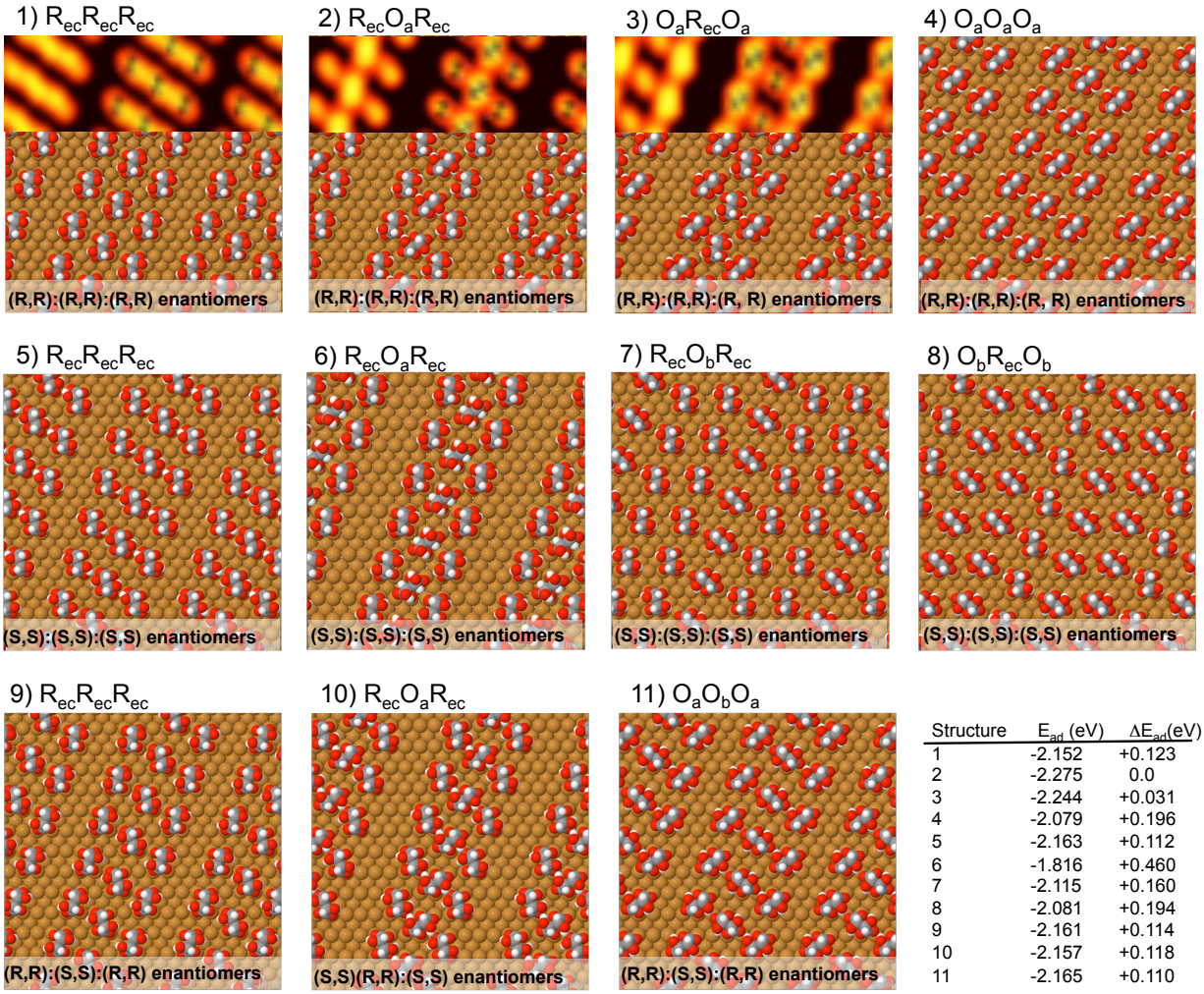


Figure 2. Calculated adsorption structures for *(R,R)*-bitartrate on Cu(110), adopting three different adsorption sites, projecting O_a, R_{ec} and O_b footprints [18]. The top two panels show the adsorption geometries with the different atoms identified as follows: Carbon: grey; Oxygen: red; Hydrogen: white. Intra-molecular H-bonds are depicted with yellow lines. The bottom panel shows the adsorption footprints created by the four O-Cu chemisorption bonds.

To distinguish between these possibilities, we turned to the additional two questions outlined above. Specifically, we considered whether the experimentally observed chiral segregation of the racemic mixture is predicted [18]. We, therefore, investigated the effect of substituting (*S,S*) and mixed *(R,R)/(S,S)* enantiomer combinations within the three minimum energy structures 1,2 and 3 calculated for *(R,R)*-bitartrate, as shown in Figure 3. The results for the originally proposed $R_{ec}R_{ec}R_{ec}$ structure 1 are particularly interesting since this assembly is, in fact, slightly more stable when *(S,S)*-bitartrate is substituted into the assembly (Structure 5). This prediction is the reverse of the experimentally observed structures for enantiopure adsorption. As a further consideration, a mixture of (R,R) and (S,S) enantiomers were incorporated within the $R_{ec}R_{ec}R_{ec}$ arrangement (structure 9). Again, the adsorption energy for mixed enantiomer occupation is slightly lower than that of the enantiopure (*R,R*) structure 1, and similar to that of the (*S,S*) structure 5. These calculations suggest that if the $R_{ec}R_{ec}R_{ec}$ footprint arrangement is adopted, the enantiomorph preference for the enantiopure systems would be flipped and no chiral segregation should be observed for the racemic system. This prediction clearly contradicts the experimental result, Figure 1, and reveals the limitations of the previously calculated structural model [14, 15] that had been accepted for bitartrate assembly on Cu(110) for over 15 years.

A similar enantiomer substitution exercise can be conducted for the other two viable footprint arrangements that can create the (1 2, -9 0) structure (Structures 2 and 3). When *(S,S)*-bitartrate occupation is considered for the $R_{ec}O_aR_{ec}$ footprint combination (Structure 6), a high penalty energy of +0.46 eV arises, largely because the *(S,S)*-enantiomer is forced to adopt the O_a footprint, rather than its preferred O_b footprint. This result illustrates the central role played by the adsorption footprint in generating an enantiospecific response during assembly, i.e. the R_{ec} footprint is insensitive to the handedness of the enantiomer, while the O_a/O_b footprints are highly enantiospecific. This strong enantiospecificity also rules out *(S,S)*-bitartrate occupation of the $O_aR_{ec}O_a$ footprint arrangement, which would be

energetically unviable. We, therefore, also examined whether it was possible for enantiopure *(S,S)*-bitartrate to generate a (1 2, -9 0) assembly by adopting a $R_{ec}O_bR_{ec}$ footprint arrangement, which allows the central *(S,S)*-bitartrate to occupy its favoured oblique footprint, but with the trimer rows still directed along the same chiral direction exhibited by the *(R,R)*-bitartrate assembly. Such an arrangement requires the trimer unit to expand in order to accommodate the *(S,S)* enantiomer in the O_b footprint, but it is possible to create an ordered a (1 2, -9 0) assembly (Structure 7). This reduces the energy penalty, but Structure 7 is still less stable than the enantiopure *(R,R)*-bitartrate Structure 2 by ~0.1 eV, which is sufficient to lead to the enantiomorph selection displayed by the enantiopure assembly and the chiral segregation of the racemic mixture at 300K. A similar exercise shows that enantiopure *(S,S)*-bitartrate occupation of the $O_bR_{ec}O_b$ trimer (Structure 8) or a *(R,R):(S,S):(R,R)* occupation of the $O_aO_bO_a$ combination (structure 11) are still less stable than the enantiopure (*(R,R)*-bitartrate structure 2.



Structure	E_{ad} (eV)	ΔE_{ad}(eV)
1	-2.152	+0.123
2	-2.275	0.0
3	-2.244	+0.031
4	-2.079	+0.196
5	-2.163	+0.112
6	-1.816	+0.460
7	-2.115	+0.160
8	-2.081	+0.194
9	-2.161	+0.114
10	-2.157	+0.118
11	-2.165	+0.110

Figure 3. Minimum energy structures calculated using DFT for various bitartrate arrangements on Cu(110) [18]. The top row displays the enantiopure *(R,R)-bitartrate* overlayers created using the combination of footprints indicated at the top of each panel. Structures 1-3 have their simulated STM images (at bias of -0.5V) superimposed on the panels. The second row shows occupation of the same footprints by the enantiopure *(S,S)-bitartrate* system. The bottom row shows occupation of the footprint template by mixed *(R,R) and (S,S)* enantiomers, with the most favourable mixed enantiomer combinations shown. The inset table shows E_{ad} of each structure and also, ΔE_{ad}, the adsorption energy relative to structure 2.

To further establish whether structure 2 is the observed experimental phase, we simulated the STM images [18] that would be generated from *(R,R)*-bitartrate assembly with structures 1, 2 and 3 (Figure 3 inset). It can be seen that only Structure 2 reproduces the experimental data, with the middle molecule in the trimer imaging brighter than the two outer ones. Therefore, overall, structure 2 in which two types of footprints are occupied, provides a better model for the chiral assembly of bitartarte on Cu(110). Crucially, the adoption of the oblique footprint within the trimer is key to

creating an energetic preference for the (R,R)-bitartrate to assemble into the (1 2, -9 0) structure, and drive chiral segregation into distinct enantiomorphs for the racemic system.

What this discussion outlines is that the nature and diversity of the individual nanoscale adsorption motifs largely determines the plethora of assemblies that could arise and, therefore, needs careful consideration. When the assembly of chiral molecules is considered, then the effect of enantiomer substitution into this large manifold of structures also needs to be considered in order to understand the chiral behaviour of the system. Finally, it should be recognised that although the diversity of viable adsorption forms add to the complexity of analysis, the flexibility to adopt a range of energetically near-equivalent conformations enables the system to respond dynamically to its environment and make local adjustments that could drive important functions, such as adaptation, host-guest accommodation, chiral recognition [20] and heterogeneous enantioselective catalysis [21, 32].

Creating Spaces within Molecular Assemblies

As important as where the molecules are in a particular assembly, it is equally important to know where the gaps and spaces are generated within an assembly, since pores and channels provide environments for important functions such as sensing, reactivity, storage, separation, selective transport, etc. Therefore, it is important to understand the factors that create spaces within molecular assemblies at surfaces. The bitartrate/Cu(110) system provides an example of surface stress induced gap creation within an assembly [14,18]. Modelling such effects has only become possible with recent advances in computational power, allowing large unit cells to be addressed. The structural model 2 of *(R,R)*-bitartrate on Cu(110) shows trimer rows separated by gaps where 3 Cu atoms in the [1 $\bar{1}$0] direction remain exposed, Figure 3. The origin of this gap arises directly from the adsorption of the molecule via the 4 oxygens of the carboxylate groups. As each carboxylate group bonds onto the close-packed row, there is a 2.5% increase in the Cu-Cu distance of the surface atoms binding to the molecule. This creates a stress within the surface, which is relieved by breaks in the assembly. Calculations [18] tracking the associated energy change as the gap between the trimer rows is increased progressively from 1 atom to 4 atoms, show that the energy of the system falls steadily until a gap of 3 atoms is reached, after which no energy advantage accrues if the gap is increased. Elsewhere in these Proceedings, papers by Nian Lin and Manfred Buck illustrate beautiful molecular assemblies leading to the creation of a wide variety of pores, while the group of de Feyter have created impressive multicomponent assemblies via a core-shell approach in which multiple components are incorporated via a host-guest interactions [34].

The Importance of Disorder in determining Ordering

The literature on molecular assembly at surfaces, by and large, focusses on the ordered arrangements that are observed. However, our work on the adsorption of scalemic mixtures of (*R,R*)- and (*S,S*)-bitartrate on Cu(110) provides an important lesson in illustrating the importance of the *disordered* phase [17]. Adsorption of the racemic 50:50 mixture produces chiral segregation as shown in Figure 4a, with enantiomorphous domains of (*R,R*)-bitartrate and (*S,S*)-bitartrate formed in equal measure at the surface. This is shown in LEED data for the system, where diffraction spots arising from each structure are present with equal intensity, Figure 4a. When enantiomerically unbalanced adsorption is carried out, the system behaviour changes radically. For example, the 60:40 mixture shows LEED spots arising from the majority species only, Figure 4b. In fact, the organisational behaviour of the two enantiomers becomes highly non-linear as shown in Figure 4c, with the organised phase for the minority species rapidly disappearing in the LEED data even for slight deviations from the racemic state. STM data for a 60:40 *RR*:*SS* system echo the LEED experiment, with large domains of the majority (*R,R*)-bitartrate visible, whereas the organised (*S,S*)-bitartrate phase is almost annihilated, appearing only in very small domains containing tens of molecules, Figure 4b. This was a very surprising result; the expectation would be that domains of each enantiomer would be observed with a 60:40 ratio. Kinetic Monte Carlo calculations undertaken by A.P.J Jansen [17] revealed that system behaviour is actually driven by the configurational entropy of the disordered 'sea' surrounding the organised domains. The greatest configurational entropy arises when the disordered phase is racemic with an equal population of the two enantiomers. This aspect has a profound influence on how the organised phases evolve. For a 60:40 adsorption mixture, there is initial preferential 2D crystallisation of the majority enantiomer to create a racemic disordered sea. Once the excess enantiomer has organised, both enantiomers start to crystallise. However, in order to maintain a racemic disordered sea via a dynamic equilibrium with the organised

phase, the domain boundaries of the segregated enantiomers must be equal. This is only possible if the majority enantiomer organises in larger domains, and the minority in smaller domains. This leads to the highly non-linear behaviour observed, with drastic deviation in the crystallisation propensity and domain size exhibited by the two enantiomers.

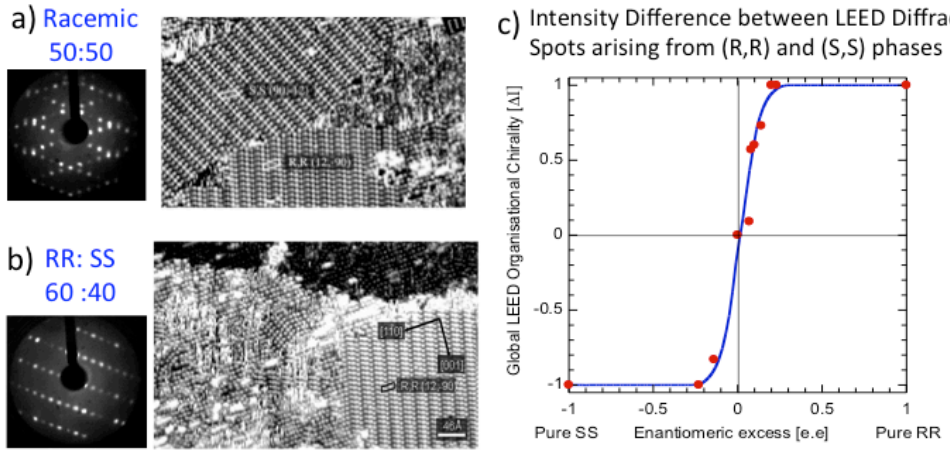


Figure 4. Non-linear behaviour in the organisation of the (*R,R*)-bitartrate and (*S,S)*-bitartrate phases on Cu(110) as the enantiomeric ratios are changed. a) LEED pattern and experimental STM image from racemic adsorption showing the organised phase from each enantiomer are formed in equal abundance; b) LEED pattern and experimental STM image from a 60:40 RR:SS mixture showing only the majority *(R,R)* enantiomer creates an organised phase, giving rise to the LEED pattern of one enantiomorphous domain only, with STM showing the the minority enantiomer is restricted to forming very small organised clusters; c) Highly non-linear behaviour in the formation of the two enantiomorphous domains (measured by LEED intensity) as the enantiomeric excess (e.e) is varied.

Amino Acid Assembly from the Single Molecule and Single Footprint Perspective: Proline on Cu(110)

The bitartrate/Cu(110) system discussed above showed the importance of considering adsorption footprints created when molecules chemisorb as a surface. Monte Carlo simulations on other systems [23] demonstrate that the evolution of molecular surface organisation and local patterns is strongly influenced by the adsorption footprints the system adopts. There is, therefore, a real need to experimentally identify the adsorption footprints that are adopted within an assembly. This is a challenging task, especially for the small molecules considered here, which image as single features in STM. The importance of footprints reaches another level when considering the adsorption of amino acids on the Cu(110) surface, as exemplified by our work on the amino acid proline and its derivatives on Cu(110) [24-28]. The amino group of proline is incorporated within a pyrrolidine ring, Figure 5iii, leading to greatly increased structural rigidity. This rigidity is instrumental in enabling identification of different adsorption footprints, since each adsorption conformation triggers a specific reorientation of the attached pyrrolidine ring, which leads to image contrasts in STM, as described below.

Adsorption of enantiopure *(S)*-proline on Cu(110) leads to a (4 × 2) assembly containing the dehydrogenated prolate species [25,26]. STM images, Figure 5ib, show the (4 × 2) unit cell contains two distinct prolate conformers, which image as bright and faint features. DFT calculations reveal that the two observed conformers arise as a direct consequence of the different adsorption footprints that are adopted by the molecule. For both conformers, the two carboxylate oxygen atoms bind to adjacent copper atoms in the close-packed row. In addition, a third bonding contact occurs with the nitrogen binding to a copper atom in the neighbouring row. It is the position of this bond that distinguishes the two conformers. When the nitrogen atom bonds to the copper atom to the left, a left-handed triangular adsorption footprint is described and the pyrrolidine ring tilts significantly away from the surface (conformer A, figure 5ii), imaging as a bright protrusion in STM. However, if the nitrogen bonds to the right, a right-handed triangular adsorption footprint is described, the ring lies flat (conformer B, figure 5ii) and images as a faint protrusion. Importantly, the left and right-handed footprints are non-superimposable mirror images in two-dimensions and are, therefore, chiral. This means that, effectively, a second level of chirality is imprinted at the surface in the form of the adsorption footprints.

The two conformers have similar adsorption energies and are both involved in the molecular assembly, and eight distinct (4 × 2) arrangements can be constructed for *(S)*-prolate on Cu(110) [25]. By calculating adsorption energies for the different assemblies and comparing their simulated STM images with the experimental data, Figure 5i, it can be determined that the enantiopure *(S)*-prolate assembly possesses an ordered heterochiral footprint arrangement, Figure

5iv. Experimentally, the STM contrast enables chirality to be tracked at the single footprint level and confirms the theoretical prediction.

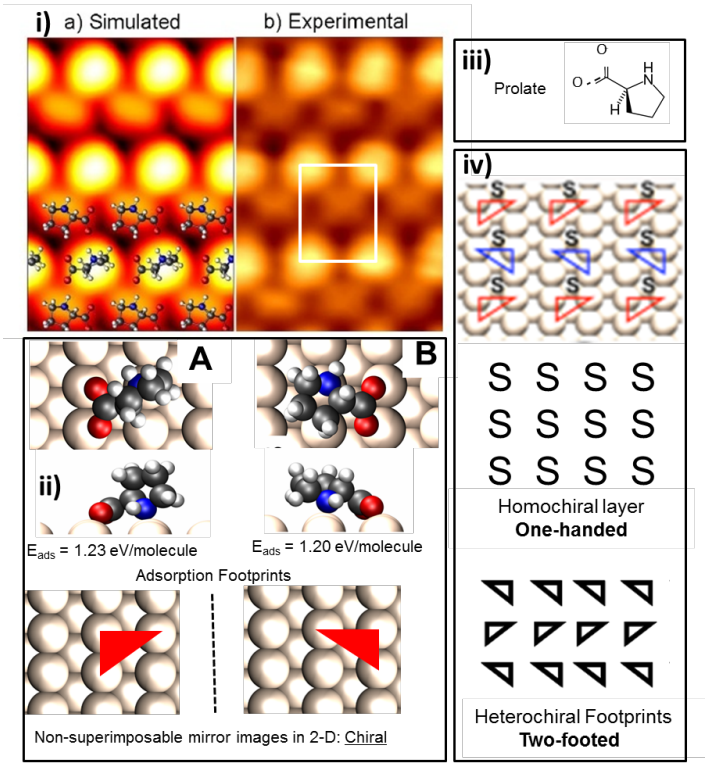


Figure 5. i) DFT simulated and experimental STM image of the (4×2) overlayer of (*S*)-prolate on Cu(110); ii) DFT calculated adsorption geometries for conformers A (bright spot in simulated image) and B (faint spot in stm images) of adsorbed prolate with views from above, the side, and in terms of their triangular adsorption footprint (atoms are depicted by: red:O, blue:N, black:C, and white:H); (iii) structure of the prolate molecule; (iv) Schematic diagram showing the (4x2) adlayer identifying the molecular and footprint chirality at each position. This information is further deconvoluted to depict the enantiomer and footprint arrays that are formed.

The molecule-by-molecule analysis of the *(S)*-prolate/Cu(110) assembly demonstrates that the molecule-surface interface now possesses two manifestations of chirality, one arising from the molecular handedness and the other from the footprint chirality, which we have chosen to label as 'footedness' [24,28,29]. In cases where single-handed molecules project footprints of one chirality only, i.e. the molecules are single-footed, one can confine the description of the interface chirality with respect to the molecular handedness only. However, if a single-handed molecule can project footprints of either chirality i.e. it is two-footed, Figure 6a, then the chiral description of the interface needs to capture this dual aspect. We have, therefore, introduced an overall chirality descriptor [29], $\boldsymbol{C}_F^H$, which describes the chiral ordering at the molecule-surface interface from both the molecular handedness (H, superscript) and the adsorption footedness (F, subscript) viewpoints, Figure 6b. Here, the classical descriptions of global chiral ordering such as enantiopure assembly, racemic conglomerate, racemic compound and random solid solution are given a label (En, Cn, Rc and Rn, respectively), as shown in Figure 6c. Thus, the *S*-prolate/Cu(110) assembly has an overall chirality of $\boldsymbol{C}_{Rc}^{En}$.

Determining the overall chirality of such an interface is particularly difficult to determine experimentally since techniques such as reflection absorption infrared spectroscopy (RAIRS), near edge x-ray absorption fine structure (NEXAFS) and photoelectron diffraction (PhD) are averaged across the whole surface and do not provide information at the single-molecule level. Here, the rigidity of the prolate and the re-orientation of the ring as the adsorption footprint is altered provides direct mapping of this aspect from the STM images.

Such analysis can be taken further by considering adsorption of the racemic mixture, where four diastereomers of handedness and footedness can arise, Figure 6a. The complexity of the overlayer that could be generated by such a system is evident from the STM image of racemic prolate on Cu(110) [26], Figure 7i. Although the arrangement of the molecules follows the (4x2) pattern observed for the enantiopure system, there appears a random distribution of bright and faint protrusions. This complex image can be deconvoluted by considering high contrast images obtained from the

enantiopure systems, Figure 7ii, which show 4 distinct features, each associated with a particular handedness-footedness diastereomer. This now provides a key to read the entire STM image, as shown in Figure 7ib-d, where a portion of the STM image is analysed in terms of chirality at the single-molecule and single-footprint level.

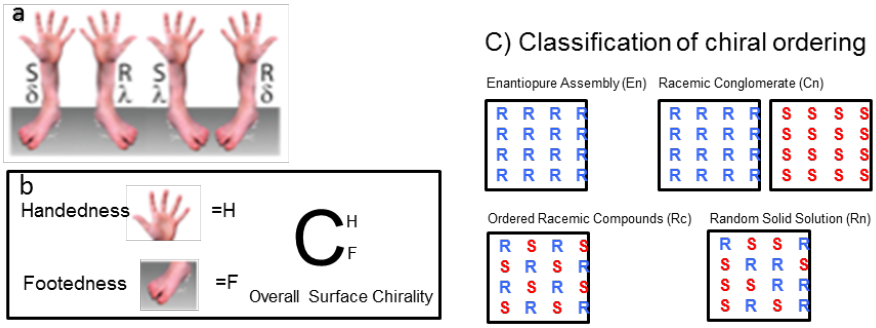


Figure 6: a) Depiction of the dual manifestation of chirality at surfaces created by both handedness (R and S) and footedness (λ and δ); b) Description of overall surface chirality as C_F^H where the organisation of handedness is depicted by (H, superscript) and of footedness by (F, subscript); c) The ordering description is based on the different classes of chiral ordering and is indicated as follows: enantiopure (En); racemic conglomerate (Cn), racemic compound (Rc) and random solid solutions (Rn).

The ensuing arrangement can be deconvoluted into separate maps of molecular chirality and maps of adsorption footprints, Figure 7id. From this, it can be seen that the enantiomer arrangement within the adlayer is entirely random, i.e. the molecular assembly is a two-dimensional random solid solution. However, the adsorption footprints adopt the strict heterochiral template that is exhibited by the enantiopure *(S)*-prolate system. Therefore, racemic prolate/Cu(110) produces an overall chirality of C_{Rc}^{Rn}. Clearly in this system, the chirality of the adsorption footprint, and not that of the molecule, drives the arrangements of both enantiopure and racemic proline at the surface.

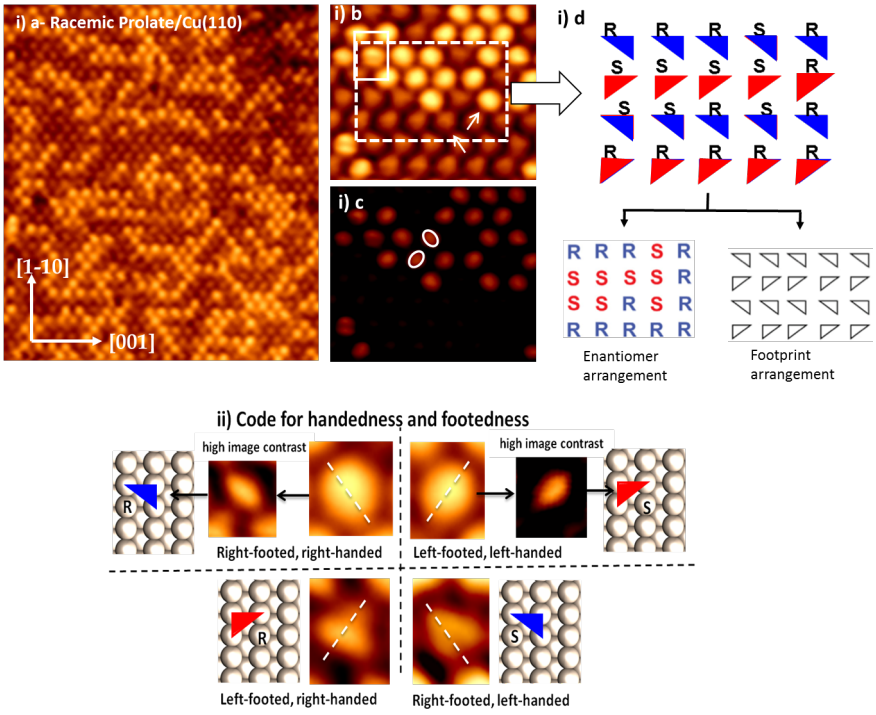


Figure 7. i)a STM image of the racemic (*R*,*S*)-prolate overlayer on Cu(110) (158×125 Å^2, *I*(*t*)=−0.62 nA, *V*(*t*)=−1238 mV); i)b and i)c High-resolution images showing the bright and faint features in the STM overlayer (44×50 Å^2, at differing image contrast; 1)d Schematic diagram identifying the molecular and footprint chirality at each position within the (4x2) adlayer contained within the outlined square. This information is further deconvoluted to give the enantiomer and footprint arrangements at the surface. ii) Code for interpreting each of the four distinct STM features in terms of molecular and footprint chirality.

The complexity of assembly exhibited in these and other amino acids on Cu(110) has recently [29] been rationalised using 3 simple rules, described below:

i. The carboxylate rule: '*The carboxylate groups of neighbouring molecules avoid repeated binding in adjacent sites*'. This rule arises because electrostatic repulsion and surface strain create a large energy penalty if carboxylate groups are placed at nearest neighbour positions along an infinitely repeating 1D chain. As a result, for periodic (3 × 2) and (4 × 2) assemblies observed for amino-acids on Cu(110), the carboxylate groups are staggered in both high symmetry directions of the surface.

ii. The footedness rule: *'The allowed ordering at both the handedness and footedness levels is determined by the chirality of the adsorption footprints that can be created at the surface and the manner in which individual enantiomers may occupy these footprints'*. The footedness rule encapsulates two aspects: (a) the generation of allowed footprint templates; and, (b) occupation of these templates by enantiomers.

iii. The interactions rule: *'Of the outputs allowed by the carboxylate and footedness rules, the overall chirality that a molecular overlayer ultimately adopts will be determined by a balance between the optimization of molecule-substrate interactions and intermolecular interactions'*. These interactions will clearly be system-specific.

For the prolate/Cu(110) system, eight organised (4x2) templates can be suggested by the carboxylate rule, of which 2 are homochiral and 6 are heterochiral [29]. DFT calculations show that the specific heterochiral (4x2) footprint arrangement adopted is one that minimises repulsive carboxylate-carboxylate interactions and maximises intermolecular hydrogen bonding [25,26]. The reason that the C_{Rc}^{En} and the C_{Rc}^{Rn} assemblies emerge for enantiopure and racemic prolate adsorption, respectively, arises from the footedness rule. Figure 8a compares how the footprints in a template would be occupied by enantiomers for a single-footed and a two-footed system. Such considerations allow a 'surface chirality generator' to be run for each template. Figure 8b shows this for the preferred heterochiral template adopted by the prolate/Cu(110) system. It can be seen that the creation of the C_{Rc}^{En} overlayer for enantiopure adsorption and the C_{Rc}^{Rn} overlayer for racemic adsorption are now self-evident. The two-footed aspect for the racemic assembly is further confirmed by calculations that show that within the adopted heterochiral footprint arrangement, the upright conformer of (R)-proline is equivalent to the flat-lying conformer of (S)-proline, with respect to the formation of hydrogen-bonded chains. Thus, each adsorption position in the footprint template can be occupied by either enantiomer resulting in a 2-D random solid solution at the surface [26].

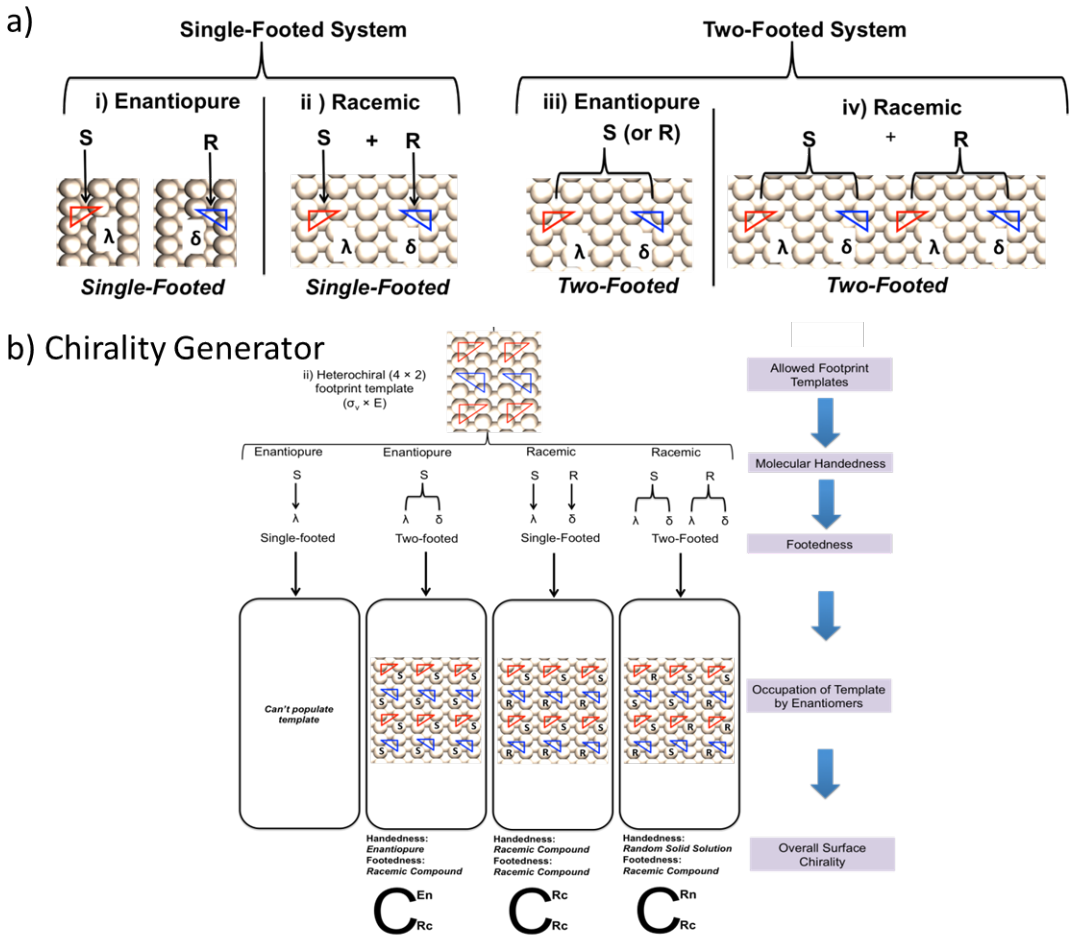


Figure 8. a) The occupation of footprints by chiral molecules in: (i) and (ii) a single footed system with the *(S)*-enantiomer placing the λ-footprint and the *(R)*-enantiomer placing the δ-footprint; (iii) and (iv) a two-footed system where each enantiomer may place either the λ or δ-footprint. (red triangles = left λ-footprint, blue triangles = right δ-footprint). The behaviour for both enantiopure and racemic systems is shown. b) Depiction of the surface chirality generator, which predicts the overall chiral ordering that would arise for the heteochiral (4x2) footprint template, as the molecular handedness and footedness of the molecules are altered.

The simple rules derived for the amino-acid assemblies on Cu(110) [29] provide a framework for capturing the allowed manifold of assemblies that could potentially arise, and are a step forward in creating design rules that target the creation of specific arrangements and outcomes. For example, one of the conclusions of our analysis is that chiral segregation is a very particular output of systems that favour homochiral templates and are single-footed [29]. Note that the bitartrate/Cu(110) system falls in this category because there is a strong energy penalty for an enantiomer to occupy a mirror footprint. Currently, the literature on amino acids shows that many prefer heterochiral footprint templates and are also two-footed. This leads to an interesting input into debates on the origin of homochirality of life [35,36], where most mirror-symmetry breaking scenarios involve chiral resolution or amplification mechanisms to create homochiral domains of monomers as a starting point for creating homochiral polymers of life. However, for almost all of these scenarios there is an equal probability of creating homochiral pools of either enantiomer, opening up equivalent pathways for creating right- and left-handed polymers. In an alternative scenario, if adsorption of a racemic mixture leads to a random solid solution of enantiomers at a surface, an inherent mechanism for mirror-symmetry breaking emerges. If the randomly organised array of enantiomers were subject to polymerisation, a large diastereomeric library of polymers would emerge, Figure 9, each different with distinct properties. If homochiral polymers provided a particularly favoured attribute, these would be selected by evolutionary pressure. However, the chance of a randomly organised assembly creating a homochiral polymer becomes progressively smaller as the length of polymer is increased. For such a scenario, the probability of creating two mirror image homochiral polymers becomes

vanishingly small as the polymer length increases beyond even a small number of units, thus providing an inherent mirror-symmetry breaking process. This idea now leads us to consider whether the surface provides an appropriate environment to covalently couple molecules into macromolecules.

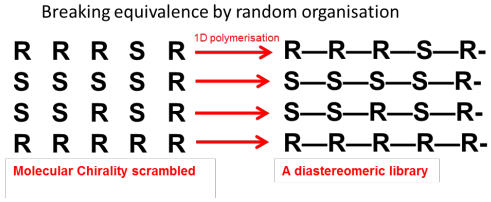


Figure 9: Schematic showing how 1D polymerisation of a random solid solution of a racemic mixture would lead to a diastereomeric library of polymers.

On-Surface Covalent Assembly of Macromolecules

The field of on-surface synthesis has moved at great pace since the first reports of macromolecular creation via direct coupling of individual molecular components at a surface [8-11]. Today, a range of approaches have been translated successfully from organic synthesis methodology to on-surface synthesis including condensation reactions [11,37-39], Ullman coupling [8,10,40-43], Glaser coupling [44-46] and dehydrogenative coupling [9,47-53] producing a variety of covalent structures at the surface, ranging from on-surface synthesis of complex molecules [41,54,55], to linear polymers [8,47,49,56], extended networks [10,45,55] and graphene nanoribbons [10,57-58]. Examples of on-surface synthesis are also illustrated in this Faraday Discussion by Lieng Chi, Nian Lin and Frederico Rosei.

In collaboration with David Amabilino, we have explored using the C-H group as the functional synthon and exploiting the surface to activate, dissociate and couple the components. Of course, C-H bond activation is well known for being challenging in organic synthesis and heterogeneous catalysis and our choice of the Cu(110) surface provided a level of reactivity, which is sufficient to enable reactions to occur, but not so great as to decompose the molecular building blocks. In addition, the anisotropy of the surface allowed regioselective behaviour to be identified. Finally, this approach opens up access to the huge range of starting materials that possess C-H groups and, also, delivers clean coupling reactions in which C-H bond scission leads to H atoms that recombine and desorb cleanly from the surface. Below, the success of this dehydrogenative coupling approach to covalently link a range of molecules at surfaces is outlined, with examples ranging from homocoupling to heterocoupling reactions involving components of different shape, size, connectivity and topology.

Covalent Coupling to create linear poly-porphyrin organometallic oligomers.

The reactions of the unfunctionalised free base H_2-porphyrin and H_2-diphenylporphyrin on Cu(110) [47, 48] provide a good example of dehydrogenative coupling, where the reaction is triggered simply by heating. This generates linear oligomers of porphyrins, aligned along the [100] direction, Figure 10a, accompanied by the evolution of hydrogen from the surface. Complex functions invariably require complex molecular assembly, so an important step in on-surface synthesis is to advance from homocoupling reactions to heterocoupling scenarios, where different building blocks can be combined. A successful example [49] of this is provided by the heterocoupling of H_2-porphyrin and Zn-diphenylporphyrin on Cu(110), to give linear co-oligomers, Figure 10b. The STM images provide a direct map of oligomer composition and sequence, with H_2-porphyrin imaging as squares and Zn-diphenylporphyrin showing serrated features arising from the two phenyl groups. It can, therefore, be seen that a random co-polymer is created when the system is heated rapidly to 650K. By altering synthesis protocols, block co-polymers in which sections of one monomer are following by sections of the other are produced Figure 10b.

STM data show that the all the homocoupled and heterocoupled oligomers possess a 10.8 Å monomer-to-monomer distance. DFT calculations [47, 48] confirm this to be consistent with the creation of organometallic C-Cu-C bonds arising from cleavage of the sp^2 C-H bond at the 3, 5 and 7 positions. The calculated model and the STM simulation for a 2 Cu-atom and a 3 Cu-atom bonded polymer are shown in Figure 10c. High-resolution STM data, e.g. Figure 10a,d, are in good agreement with the theoretical images, showing the expected core-to-core distance of 10.8 Å, with the

organometallic coupling Cu atoms imaging bright. It is noted that organometallic polymers are an important class of hybrid materials, combining the properties of organic systems with those of metals, with applications in molecular wires, photoelectronics and sensors.

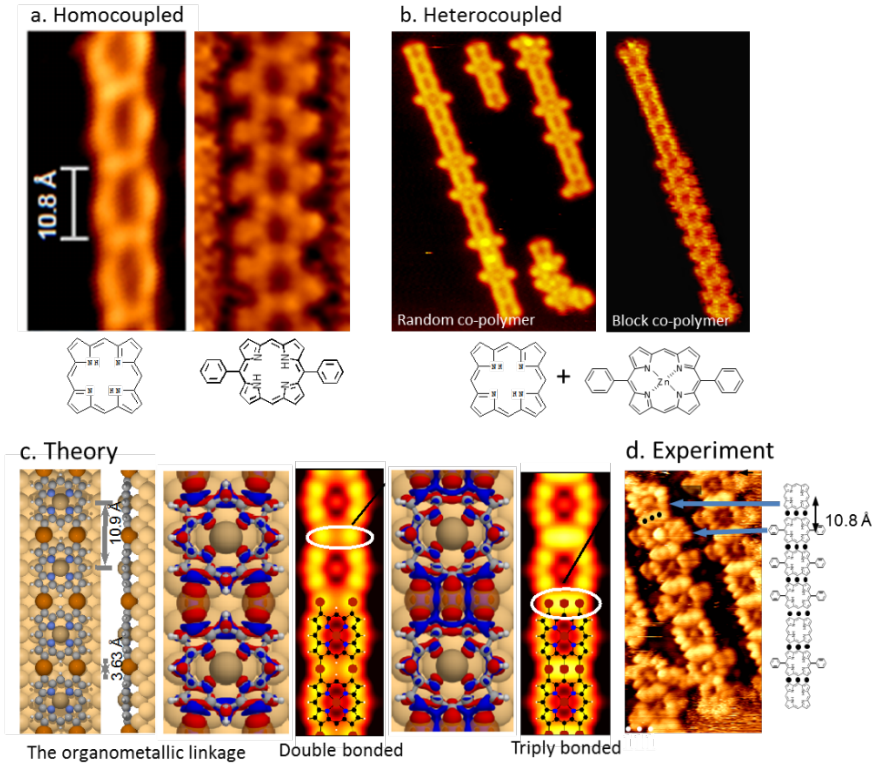


Figure 10. Covalently bonded linear oligomers created from the reaction of H_2-porphyrin, H_2-diphenylporphyrin and Zn(II) diphenylporphyrin. a) Homocoupled structures left, I_t = 0.31 nA, V_t = 0.30 V and right, I_t = 0.41 nA, V_t = -1.68 V; b) Heterocoupled structures, left: 100 x 190 A, I_t = 0.13 nA, V_t = -1.68 V, and right: 65 x 170 A, I_t = 0.14 nA, V_t = -0.57 V; c) Theoretically calculated polymer structures on Cu(110) and the corresponding simulated STM images showing the doubly and triply bonded Cu-porphyrins chains and the charge accumulation (red) and depletion (blue) that occurs when the polymer is formed. d) High resolution experimental STM image (35 x 70 A, I_t = 0.17 nA, V_t = -0.83 V) showing a heterocoupled oligomer formed by reaction of H_2-porphyrin and Zn(II)-diphenylporphyrin on Cu(110), where the chemical nature of each individual monomer and the number of interconnecting C-Cu-C bonds can be identified, as shown in the structural model.

This work reveals a number of important factors: (i) the C-H group can certainly be activated at a Cu surface, leading to C-H bond scission; (ii) the reaction involves recombinative H_2 desorption, indicating clean synthesis that does not leave side-products at the surface; (iii) organometallic coupling occurs at the surface, leading to $[porphyrin\text{-}Cu]_n$ organocopper oligomers, a reaction which has no counterpart in general organometallic synthesis, i.e., the surface creates a unique reaction environment to create new types of functional matter; (iv) directional macromolecules aligned along a specific surface direction are synthesised, despite the absence of functional groups to stereodirect the reaction.

Control of product directionality at a surface will be important in future molecular devices and DFT modelling by Hanke et al [48] of the H_2-porphyrin/Cu(110) system shows that each additional organometallic bond formed along the [001] direction, leads to an energetically more stable structure, with the most stable structures corresponding to the triply-coupled unit with Cu insertion at the 1,3,5 positions of the porphyrin, or the doubly-coupled unit with C-Cu-C coupling at the 1 and 5 position, Figure 11. An analysis of the product configurations reveals that the commensurability of the oligomer structure with the underlying surface structure along the [001] direction is also particularly favourable, accommodating the coupling Cu atoms on preferred 4-fold sites of the underlying surface as 'added rows' and causing very little distortion of the porphyrin core. However, coupling in the orthogonal [1 $\bar{1}$0] direction is not favourable because the product dimensions are less compatible with surface geometry, leading to distortion and buckling of the porphyrin ring which, in turn, decreases the bonding interaction. An energy advantage only appears when the triply

bonded or the doubly-bonded offset units are created, Figure 11, however, both still compare unfavourably with products aligned In the [001] direction.

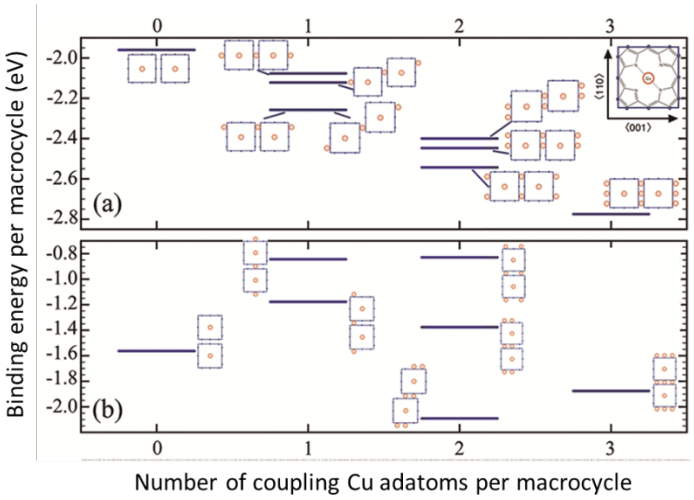


Figure 11. Binding energies per porphyrin core for covalently bonded structures aligned along the [001] direction (top panel) and along the [110] direction (bottom panel) of the Cu(110) surface. Reproduced with permission from *ACS Nano*, 2011, **5**(11), 9093 [ref 48]. Copyright (2011) American Chemical Society.

Heterocoupling of unlike molecules.

Coupling of two very different molecules is demonstrated when H_2-porphyrin and pentacene were co-adsorbed on Cu(110) and reacted together by heating to 650K [49]. STM data show a diverse set of products is synthesised at the surface via modular construction where dimer and monomer pentacene bind to the 1D porphyrin organometallic wires, Figure 12a. Inter-module connections occur either at the side of the oligoporphyrin long edge to produce simple and complex key topologies or at the porphyrin chain ends to give capped ladder topologies. When pentacene caps the porphyrin oligomers, it acts as a terminating group as shown in Figures 12b,c. Bright regions imaged at the inter-module boundaries are indicative of copper atoms linking the two covalent modules together. The STM image in Figure 12a also shows discrete homocoupled pentacene dimers, where dehydrogenation along one length of each pentacene enables 5 organometallic linkages to be formed, which are accommodated as a local 'added row' structure. This aligns the pentacene long-axis parallel to the Cu(110) rows, which agrees with DFT calculations and STM observation, Figure 12d.

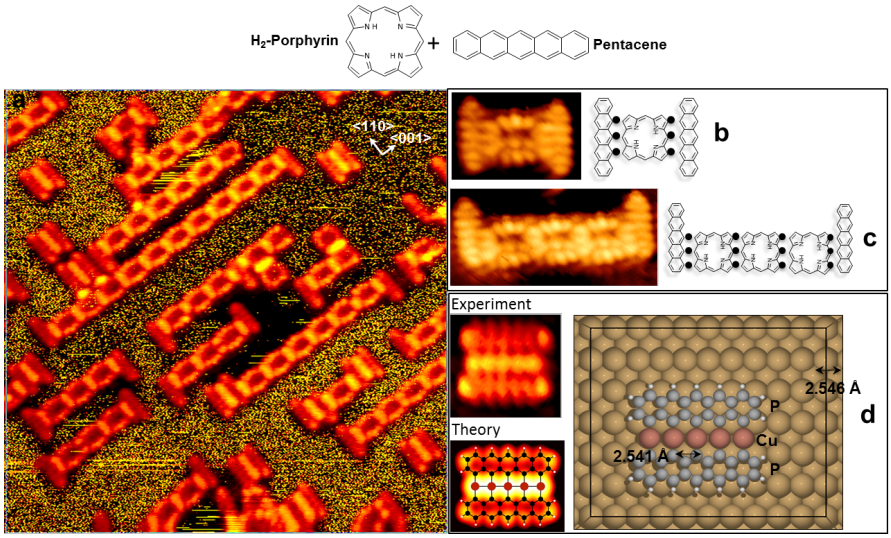


Figure 12. a) STM image (150 x150 Å^2, I_t = 1.35 nA, V_t = 0.47 V) showing the diversity of macromolecules formed by coupling reactions of H_2-porphyrin and pentacene on Cu(110) [49]; b) and c) High resolution STM showing two types of pentacene-porphyrin products, alongside chemical structural models where the black atoms represent the coupling Cu adatoms; d) Experimentally measured and theoretically calculated STM image of a covalently bonded pentacene dimer (I_t = 0.19 nA, V_t = -0.034 V, theory V_t = -1.34 V).

Covalent Coupling to create C-C coupled porphyrins.

Dehydrogenative coupling can also create C-C covalent linkages, first reported for tetra(mesityl)porphyrin on Cu(110) [9]. This reaction is particularly interesting because the 4-methyl groups act as unique connection points, displaying highly regioselective C-C bond creation. This system has recently been modelled in great detail by state-of-the-art density functional theory (DFT) and Nudged Elastic Band (NEB) calculations, to identify the thermodynamic and kinetic factors that determine the unique selectivity in covalent bonding [53]. This study deconvoluted the hierarchy of C-H bond strengths within the molecule, and demonstrated how the adsorption site and adsorbate conformation influence this hierarchy, leading to the selectivity in C-H activation and de-hydrogenation propensity at different sites on the molecule. In addition, thermodynamic and kinetic effects, including diffusion anisotropies and product accommodation at the surface determine which intermolecular C-C coupling reactions and coupled products are favoured at the surface. These insights are invaluable in unravelling the detail of the covalent assembly at surfaces and a step towards knowledge-based synthetic rules. Finally, the diversity of molecules that can be homocoupled and heterocoupled at the Cu(110) surface is shown schematically in Figure 13. This demonstrates that the C-H synthon provides a highly generic route for clean, on-surface synthesis. This approach is particularly useful for parallel synthesis of multivariate libraries of covalent matter at a surface for systems where the design parameters to deliver specific functions are unknown (which, at present, is generally the case).

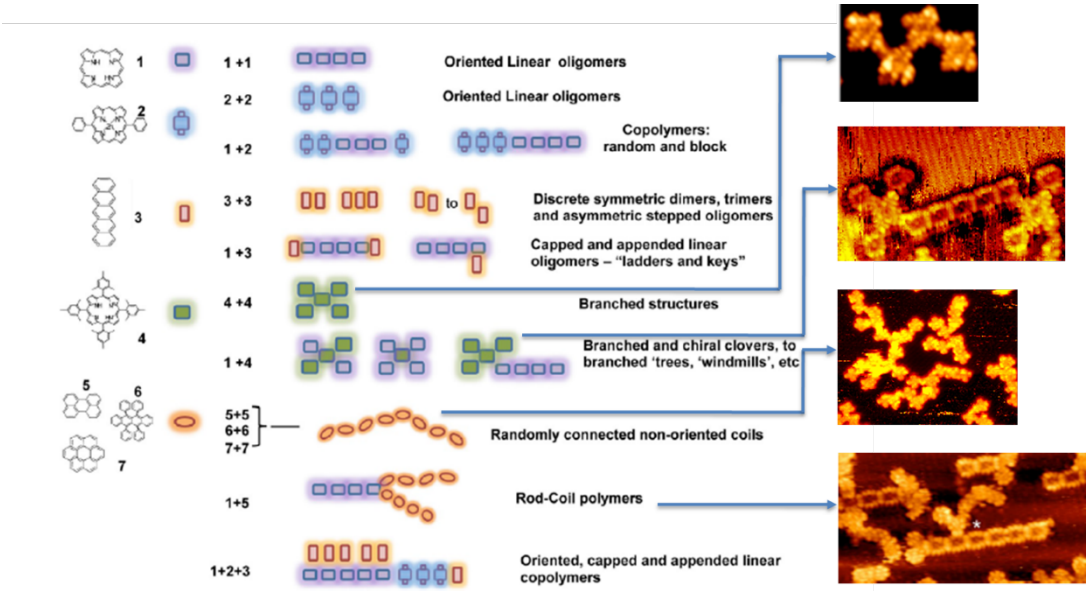


Figure 13. Summary and overview (with selected STM images) of successfully covalently coupled molecules on Cu(110) [49]. 2H-porphyrin **1**, Zn(II)-diphenylporphyrin **2**, pentacene **3**, tetramesitylporphyrin **4**, perylene **5**, rubrene **6** and coronene **7**. Reproduced with permission from *ACS Nano*, 2014, **8**(9), 8856 [ref 49]. Copyright (2014) American Chemical Society.

Creating Functional Molecular Systems

We are still at the earliest stages of creating 'bottom-up' functional molecular devices at surfaces. Nanopore sequencing technology provides a glimpse of the future potential [59,60], however this successful approach requires the use of a biological entity, e.g. the α-hemolysin pore, as the chief operational component. Elsewhere, molecular electronics has utilised single molecule components and self-assembled molecular layers to deliver function [61,62]. In terms of de novo design of molecular devices, the field of molecular machines certainly stands out, with the 2016 award of the Nobel Prize in Chemistry to Sauvage, Stoddard and Feringa. An important achievement is to translate such entities at a surface and this has been successfully demonstrated recently for selected systems [63-65].

A future direction is 'bottom-up' assembly of simple components at a surface to create a device. The example below [66]

shows an approach we created in collaboration with David Amabilino and Lluisa Pérez-Garcia, with insights into system behaviour modelled in collaboration with Lev Kantorovitch. The aim of the study was to create a surface system in which a simple molecular walker is confined to walk along a specified path between two locations. Such a system would be a mimic of the rotaxane system [67], where a macrocycle shuttles along an axis confined between two stations, or the biological linear protein motor kinesin on microtubule filaments [68,69]. An important aim was to achieve room temperature operation. Therefore, a divalent bis(imidazolyl) molecule, Figure 14a, was synthesised, with the imidazolyl functionality providing attachment 'feet' that would chemically bond and detach from the close-packed Cu(110) rows, which would act as a linear track for motion. To create barriers on the track, porphyrin oligomers were synthesised directly on the surface in an orthogonal orientation to the track would, thus creating a 'fence' to confine the motion. Figure 14b depicts a schematic of the proposed system while Figure 14c shows the STM image of the system at room temperature. The motion of the molecular walker is imaged as intensity streaks, which are clearly aligned along the [1 $\bar{1}$0] surface axis showing directional motion. It is also clear that the motion is confined between the immobile porphyrin fences. At low temperature, the motion of the walker molecule can be frozen and each individual molecule can be imaged, Figure 14d. DFT and NEB calculations show that the walker moves by attaching and detaching its imidazolyl feet in an inchworm fashion along the [1 $\bar{1}$0] direction, with directionality of motion arising from specific chemisorption interaction with the surface and from highly anisotropic diffusion barriers, which make motion in the orthogonal [001] direction energetically expensive. Given the ability of porphyrins to heterocouple with a number of different molecules, Figures 10 and 12, the opportunity now opens up to create non-equivalent "stations" which could incorporate recognition or stimuli-triggered functions that could initiate responsive motions.

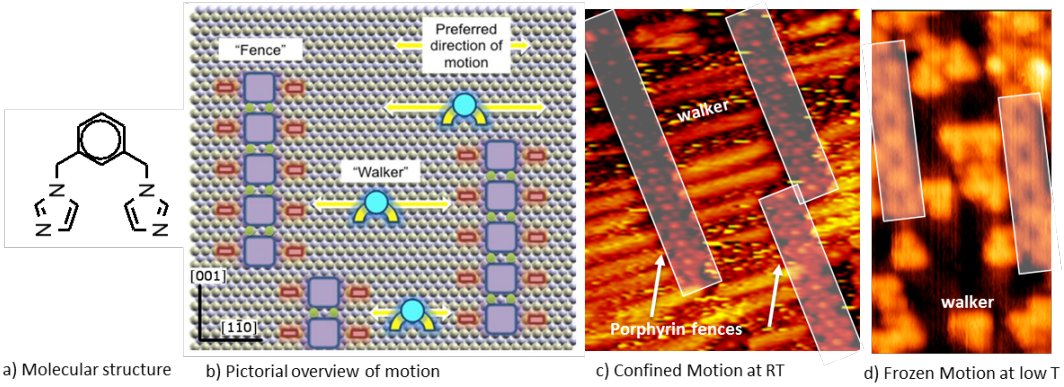


Figure 14. a) The molecule structure of the 'walker' (1,3-bis(imidazol-1-ylmethyl)benzene). b) Pictorial overview of the two-component molecular system designed to deliver confined, directional motion of a walker molecule between covalent porphyrin oligomer fences. c) STM image of the system at 300K showing diffusion streaks associated with the motion of the walker molecule along the [1 $\bar{1}$0] direction, confined between the porphyrin fences, which are highlighted by rectangles in the images. d) STM image of the system at 100 K when the diffusion of the walker is frozen and each molecule is observed as a triangular feature.

CONCLUSIONS

Molecular assembly at surfaces and interfaces has made rapid advances, delivering ever more complex systems and utilising an increasing toolbox of molecules and coupling approaches. Concurrently, the deployment of high resolution imaging, spectroscopic and diffraction techniques and advanced theoretical modelling is yielding increasingly detailed insights. What is clear from this burgeoning body of work is the molecule-surface system possesses great diversity and polymorphism, arising from the manifold of energetically viable structures the system can adopt. This aspect presents a real challenge to both the experimentalist and theoretician. Generally, the global assembly that is adopted can be traced back to effects that are already encoded into the system at the single molecule level. Here the local flexibility and response of the molecule with respect to the balance of molecule-surface and molecule-molecule interactions plays a critical role. A corollary of this behaviour is that such systems possess responsivity, adaptability and tunability. However, knowledge-based engineering of such finely balanced systems will require multiscale analysis of the system, from the nanoscale to the macroscale. For example, little is understood of the influence and balance of the many interactions that drive molecular assembly at surfaces, such as molecule-molecule interactions encompassing non-

directional van der Waals and directional H-bonding interactions, alongside strong electrostatic and covalent interactions. Equally, the role of surface-molecule interactions, ranging from van der Waals to strong electrostatic charge transfer, direct metal-molecule bonding and surface reconstructions need to be understood. There is also ample evidence that the kinetics of self-assembly is a major factor in determining the final structure, often driving polymorphic behaviour and leading to widely varied outcomes, depending on the conditions of formation. Thus experimental and theoretical approaches need to be progressed in order to capture system kinetics, including diffusion barriers and transition states. This will require time-resolved techniques and computational methods that can encompass reasonable time domains. Finally, it is pertinent to note that it took many decades of effort to create the synthetic principles of organic chemistry. It will also take our field an equivalent time to lay down the design principles of molecular assembly at surfaces.

Acknowledgements

This article draws on the work of many fine researchers and collaborators it has been my pleasure to work with. Their work has been referenced within the manuscript and particular acknowledgement goes to members of my group: Matthew Forster, Sam Haq, Mendel In't Veld, Ning Liu, Andrew Mark and Bart Wit; and to my collaborators and colleagues: David Amabilino, George Darling, Matthew Dyer, Felix Hanke, Tonek Jansen, Andreas Floris, Andrew Hodgson, Lev Kantorovitch, Mats Persson and Lluisa Pérez-Garcia. I would like to thank them all for their contributions as we made this scientific journey.

I also thank EPSRC (EP/J019364/1, EP/K039687/1, EP/F00981X/1) and EU FP7 projects RESOLVE (NMP4-SL-2008-214340) and SMALL (PITN-GA-2009-238804) programmes for supporting this research.

Notes and references

1 D.B. Amabilino, Supramolecular Chemistry at Surfaces, 2015, DOI: 10.1039/9781782622161-FP001, eISBN:978-1-78262-216-1.

2 S.M. Barlow and R. Raval, *Surface Science Reports*, 2003, **50**, 201.

3 J. A. A. W. Elemans, I. De Cat, H. Xu, S. De Feyter, *Chem. Soc. Rev.* 2009, **38**, 722.

4 T. Balandina, K. Tahara, N. Sändig, M. O. Blunt, J. Adisoejoso, S. Lei, F. Zerbetto, Y. Tobe, and S. De Feyter, *ACS Nano*, 2012, **6** (9), 8381.

5 L. Bartels**,** *Nature Chemistry,* 2010, **2**, 87.

6 F. Kling, R. Bechstein, P. Rahe and A. Kuhnle, Self-Assembly of Organic Molecules on Insulating Surfaces, Springer Series, Heidelberg, 2015.

7 G.R. Desiraju**,** *J. American Chem. Soc*., 2013, ***135*** (27), 9952.

8 L. Grill, M. Dyer, L. Lafferentz, M. Persson, M.V. Peters, and S. Hecht, *Nature Nanotech.,* 2007**, 2**, 687.

9 M. In't Veld, P. Iavicoli, S. Haq, D.B. Amabilino and R. Raval, *Chem. Commun.* **2008,** 1536.

10 J. Cai, P. Ruffieux, R. Jaafar, M. Bieri, T. Braun, S. Blankenburg, M. Muoth, A.P. Seitsonen, M. Saleh, X.Feng, K. Müllen, R. Fasel, *Nature* 2010**, 466**, 470.

11 M. Matena, T. Riehm, M. Stöhr, T.A. Jung, A. Thomas, L.H. Gade, *Angew. Chem. Int. Ed.* 2008, **47**, 2414.

12 M.O. Lorenzo, C.J. Baddeley, C. Muryn, R. Raval, *Nature* 2000, ***404*** (6776), 376.

13 M.O. Lorenzo, S. Haq, T. Bertrams, P. Murray, R. Raval, C.J. Baddeley, *Journal of Physical Chemistry B* 1999**,** ***103*** (48), 10661.

14 L. Barbosa, P. Sautet, *J. American Chem. Soc*., 2001, ***123*** (27), 6639.

15 C.G.M. Hermse, A.P. van Bavel, A.P. J.Jansen, L. Barbosa, P. Sautet, R. van Santen, *Journal of Physical Chemistry B,* 2004, ***108*** (30), 11035.

16 R. Fasel, J. Wider, C. Quitmann, K-H. Ernst, T. Greber, *Angew. Chem. Int. Ed*., 2004, **116**, 2913.

17 S. Haq, N. Liu, V. Humblot, A. P.J. Jansen and R. Raval, *Nature Chemistry*, 2009, **1**, 409.

18 G. Darling, M. Forster, C. Lin, N. Liu, R. Raval and A. Hodgson, *Phys. Chem. Chem. Phys*., 2017, **19**, 7617.

19 The overlayer unit mesh is given in real space matrix notation as follows and quoted in the text as ($_{m11\ m12}$, $_{m21\ m22}$):
$$\begin{pmatrix}\mathbf{a}_o\\ \mathbf{b}_o\end{pmatrix}=\begin{pmatrix}m_{11}m_{12}\\ m_{21}m_{22}\end{pmatrix}\begin{pmatrix}\mathbf{a}_s\\ \mathbf{b}_s\end{pmatrix}$$
where $_{ao}$, $_{bo}$ are the overlayer net vectors and $_{as}$, $_{bs}$ are the underlying metal surface mesh vectors.

20 N. Liu, S. Haq, G. R. Darling, R. Raval, *Angewandte Chemie-International Edition,* 2007, ***46*** (40), 7613.

21 A. J. Gellman, W. T. Tysoe, F. Zaera, *Catalysis Letters*, 2015, **145** (1), 220.

22 M Ortega-Lorenzo, V. Humblot, P.Murray, C.J.Baddeley, S.Haq and R.Raval, J. Catalysis, 2002, **205,** 123.

23 P. Szabelski, *Chem. Eur. J.* 2008, **14**, 8312.

24 M. Forster, M.S. Dyer, M. Persson, R. Raval, *J. American Chem. Soc*., 2011, **133** (40), 15992.

25 M. Forster, M.S. Dyer, M. Persson, R. Raval, *J. American Chem. Soc*., 2009, **131** (29), 10173.

26 M. Forster, M.S. Dyer, M. Persson, R. Raval, *Angewandte Chemie-International Edition*, 2010, **49** (13), 2344.

27 M. Forster, M. Dyer, S.D. Barrett, M. Persson and R. Raval, *J. Phys. Chem. C.*, 2011, **115,** 1180.

28 A.G. Mark, M. Forster and R. Raval, *ChemPhysChem*, 2011, **12**, 1474.

29 M. Forster and R. Raval, *Chemical Communications*, 2016, **52**, 14075.

30 R.Fasel, M. Parschau, K.-H. Ernst, *Nature*, 2006, **439**,449.

31 K.Tahara, H. Yamaga, E. Ghijsens, K. Inukai, J. Adisoejoso, M.O. Blunt, S. De Feyter and Y.Tobe, *Nat. Chem*., **2011**, 3, 714

32 J.-C. Lemay, Y. Dong, M. N. Groves, V. Demers-Carpentier, G. Goubert, R. Lafleur-Lambert, J. Boukouvalas, B. Hammer, P. H. McBreen, *Science*, 2016, **646**, 13.

33 M. Lingenfelder, G. Tomba, G. Constantini, L. C. Ciacchi, A. De Vita, K. Kern, *Angew. Chem. Int. Ed*. 2007, **46**, 4492.

34 R G. Velpula, T. Takeda, J. Adisoejoso, K. Inukai, K. Tahara, K. S. Mali, Y. Tobe and S. De Feyter, Chem. Commun., 2017, 53, 1108

35 M. Avalos, R. Babiano, P. Cintas, J. L. Jimenez, J. C. Palacios, *Orig. Life Evol. Biosph.* 2004, **34**, 391.

36 I. Weissbuch, G. Bolbach, L. Leiserowitz, M. Lahav, *Orig. Life. Evol. Biosph*., 2004, **34**, 79.

37 M. Treier, N. V. Richardson, and R. Fasel, *J. American Chem. Soc*., 2008, **130**(43), 14054.

38 A.C. Marele, R. Mas-Ballesté, L. Terracciano, J. Rodríguez-Fernández, I. Berlanga, S.S. Alexandre, R. Otero, J.M. Gallego, F. Zamora, and J.M. Gómez-Rodríguez. *Chemical Communications*, 2012, **48**(54), 6779.

39 J.F. Dienstmaier, D.D. Medina, M. Dogru, P. Knochel, T. Bein, W.M. Heckl, and M. Lackinger., *ACS Nano*, 2012, **6**(8), 7234.

40 L. Lafferentz, V. Eberhardt, C. Dri, C. Africh, G. Comelli, F. Esch, S. Hecht and L. Grill, *Nature Chemistry*, 2012, **4**(3), 215.

41 Q. Fan, J. Dai, T. Wang, J. Kuttner, G. Hilt, J.M. Gottfried, and J. Zhu, *ACS Nano*, 2016, **10**, 3747.

42 M. Kittelmann, P. Rahe, M. Nimmrich, C. M. Hauke, A. Gourdon, A. Kühnle., *ACS Nano*, 2011, **5**(10), 8420.

43 R. Gutzler, L. Cardenas, J. Lipton-Duffin, M. el Garah, L. E. Dinca, C. E. Szakacs, C. Fu, M. Gallagher, M. Vondrácek, M. Rybachuk, D. F. Perepichka, and F. Rosei. *Nanoscale*, 2014, **6**(5), 2660.

44 Y.-Q. Zhang, N. Kepcija, M. Kleinschrodt, K. Diller, S. Fischer, A.C. Papageorgiou, F. Allegretti, J. Björk, S. Klyatskaya, F. Klappenberger, M. Ruben, and J.V. Barth. *Nature Communications*, 2012, **3**, 1286.

45 H.-Y. Gao, H. Wagner, D. Zhong, J.-H. Franke, A. Studer, and H. Fuchs. *Angewandte Chemie - International Edition*, 2013, **52**(14), 4024.

46 H. Zhou, J. Liu, S. Du, L. Zhang, G. Li, Y. Zhang, B.Z. Tang, and H.J. Gao, *J. American Chem. Soc*., 2014, **136**(15), 5567.

47 S. Haq, F. Hanke, M.S. Dyer, M. Persson, P. Iavicoli, D.B. Amabilino, and R. Raval. *J. American Chem. Soc*., 2011, **133**(31), 12031.

48 F. Hanke, S. Haq, R. Raval, and M. Persson, *ACS Nano*, 2011, **5**(11), 9093.

49 S. Haq, F. Hanke, J. Sharp, M. Persson, D. B. Amabilino, and R. Raval. *ACS Nano*, 2014, **8**(9), 8856.

50 Q. Sun, C. Zhang, H. Kong, Q. Tan, and W. Xu, *Chemical Communications*, 2014, **50**, 11825.

51 A. Wiengarten, K. Seufert, W. Auwärter, D. Ecija, K. Diller, F. Allegretti, F. Bischoff, S. Fischer, D.A. Duncan, A.C. Papageorgiou, F. Klappenberger, R.G. Acres, T.H. Ngo, and J.V. Barth. *J. American Chem. Soc*., 2014, **136**(26), 9346.

52 D. Zhong, J.-H. Franke, S. K. Podiyanachari, T. Blomker, H. Zhang, G. Kehr, G. Erker, H. Fuchs, and L. Chi, *Science*, 2011, **334**(6053), 213.

53 A. Floris, S. Haq, M. In't Veld, D.B. Amabilino, R. Raval, and L. Kantorovich, *J. American Chem. Soc*., 2016, **138**(18), 5837.

54 C.J. Villagómez, T. Sasaki, J.M. Tour, and L. Grill, *J. American Chem. Soc*., 2010, **132**(47), 16848.

55 M. Abel, S. Clair, O. Ouurdjini, M. Mossayon, L. Porte, J. American Chem. Soc., 2011, 133, 11203.

56 J. A. Lipton-Duffin, J. A. Miwa, M. Kondratenko, F. Cicoira, B. G. Sumpter, V. Meunier, D. F. Perepichka, and F. Rosei, *Proc. Natl. Acad. Sci. USA* 2010, **107**, 11200.

57 J. Björk, S. Stafström, and F. Hanke, *J. American Chem. Soc*., 2011, **133**(38), 14884.

58 Y.-C. Chen, D.G. de Oteyza, Z. Pedramrazi, C. Chen, F.R. Fischer, and M.F. Crommie, *ACS Nano,* 2013, **7**, 6123.

59 D. Branton et. al., Nature Biotech., 2008, **26**, 1146.

60 S. Mitsuhashi, S. Nakagawa, M. Ueda, T. Imanishi, H. Mitsuhashi, https://doi.org/10.1101/157040

61 J.R Heath, *Annu. Rev. Mater. Res*., 2009, **39**,1.

62 X. Chen, M. Roemer, L. Yuan, W. Du, D. Thompson, E. del Barco, C.A. Nijhuis, *Nature Nanotechnology*, 2017 doi:10.1038/nnano.2017.110

63 Y. Shirai, A.J. Osgood, Y. Zhao, K.F. Kelly, J.M. Tour, *Nano Lett.* 2005, **5**, 2330.

64 T. Kudernac, N. Ruangsupapichat, M. Parschau, B. Macia, N. Katsonis, S.R. Harutyunyan, K.H. Ernst, B.L. Feringa, *Nature*, 2011, **479**, 208.

65 J. Bern, D.A. Leigh, M. Lubomska, S.M. Mendoza, E.M. Perez, P. Rudolf, G. Teobaldi, F. Zerbetto, *Nat. Mater*., 2005, **4**, 704.

66 S. Haq, B. Wit, H. Sang, A. Floris, Y. Wang, J. Wang, L. Pérez-García, L. Kantorovitch, D.B. Amabilino, R. Raval, *Angewandte Chemie. Int. Ed.,* 2015, **54**, 7101.

67 J.A. Bravo, F.M. Raymo, J.F. Stoddart, A.J.P. White, D. J. Williams, *Eur. J. Org. Chem.* 1998, **11**, 2565.

68 A. Yildiz , M. Tomishige, R.D. Vale, P.R. Selvin, *Science*, 2004, **303** (5658), 676.

69 V. Balzani, A. Credi, M. Venturi, *ChemPhysChem***,** 2008, **9**, 202.